PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Are self-reported unhealthy food choices associated with an increased risk of breast cancer: prospective cohort study using the British Food Standards Agency Nutrient Profiling System
AUTHORS	Deschasaux, Mélanie; Julia, Chantal; Kesse-Guyot, Emmanuelle; Lécuyer, Lucie; Adriouch, Solia; Méjean, Caroline; Ducrot, Pauline; Péneau, Sandrine; Latino-Martel, Paule; Fezeu, Léopold; Fassier, Philippine; Hercberg, Serge; Touvier, Mathilde

VERSION 1 - REVIEW

REVIEWER	Holly R. Harris Assistant Member, Fred Hutchinson Cancer Research Center, United States
REVIEW RETURNED	03-Oct-2016

GENERAL COMMENTS	This manuscript presents the evaluation of the FSA NPS DI, which measures the nutritional quality of the diet, and breast cancer risk. A thorough discussion of the use of different food indexes is included in the discussion section. Overall the study was well-designed and the paper is well-written, however there are a few ways in which the manuscript could be improved. Methods: How was death ascertained? Please describe. Were there any breast cancer cases in age <35 years? How does including women <35 years in the analysis influence the results? If women turned 35 during the course of the study were they included in the analysis? Results: The authors describe the results as being similar in premenopausal and postmenopausal women but the HR for premenopausal women is much stronger (HR 2.46 vs 1.25) and the result for postmenopausal women (the group with the most power) is not significant. The difference in the point estimates should be more accurately described. In addition, a more formal test for heterogeneity between these estimates should be conducted and presented in the results. Results are not presented by hormone receptor status due to lack of power. However, given the differing results by menopausal status is would also be interesting to see the results by hormone receptor status, at least by ER+/PR+ and ER-/PR-, with the caveat of limited power. Discussion:

	A discussion of stronger results were observed for premenopausal compared to postmenopausal women is needed. The median follow-up time in this study is 4 years which is shorter than many previous studies that have examined dietary factors and breast cancer. The paper could be improved with a discussion of how the timing of the diet in relation to breast cancer diagnosis could influence the results. Given that in this study dietary intake is relatively close to breast cancer diagnosis and cancer may take years to develop does this indicate more a promoter effect (in contrast to being an initiator) of diet on breast cancer? How does this compare to what has been observed with previous studies that have examined dietary indices and breast cancer risk? Table 1: A few key foods/foods groups/etc could be included in Table 1 (i.e. servings of red or processed meat, fiber, vegetable intake, etc). BMI does not yet exhibit a clear trend (Q5 has the second highest mean BMI with Q1 the highest) yet the p for trend is very significant. How is this explained? Table 2: It would be helpful to presented the age-adjusted models as well.

REVIEWER	Kelly Hirko Michigan State University College of Human Medicine USA
REVIEW RETURNED	12-Oct-2016

GENERAL COMMENTS	1. It would be interesting to evaluate which component(s) of the FSA NPS score were most strongly associated with breast cancer risk 2. Although authors likely don't have power to look at ER- tumors, it might be interesting to conduct sensitivity analyses only among the ER+ to see if the associations appear weaker, given the scientific evidence that nutritional factors may be more important in ER- breast cancer. 3. What do authors mean by 'women with a null follow-up'? Please clarify 4. How many women in the NutriNet-Sante cohort were excluded from this study because they didn't provide at least 3 24-h dietary records? Did those excluded differ from included participants based on demographic characteristics related to breast cancer risk? 5. Would appreciate an explanation on how authors decided on confounders to include in model and whether results differed in age-adjusted vs. multivariable models. 6. Might be interesting to see if associations differ by BMI 7. Discussion section could benefit from a more thorough explanation of findings from prior studies of other dietary scores (AHEI, aMED, DASH) and breast cancer risk; specifically focusing on stronger findings with ER- tumors. The results from the current study are not necessarily in line with results from many studies

	which suggest null relations between diet quality and breast cancer risk overall, although inverse associations have been observed for ER- subtypes. Authors should address why they believe that the diet score in this study was strongly associated with overall breast cancer risk in this study.

REVIEWER	Kelly Hirko Michigan State University College of Human Medicine USA
REVIEW RETURNED	12-Oct-2016

GENERAL COMMENTS	1. It would be interesting to evaluate which component(s) of the FSA NPS score were most strongly associated with breast cancer risk 2. Although authors likely don't have power to look at ER- tumors, it might be interesting to conduct sensitivity analyses only among the ER+ to see if the associations appear weaker, given the scientific evidence that nutritional factors may be more important in ER- breast cancer. 3. What do authors mean by 'women with a null follow-up'? Please clarify 4. How many women in the NutriNet-Sante cohort were excluded from this study because they didn't provide at least 3 24-h dietary records? Did those excluded differ from included participants based on demographic characteristics related to breast cancer risk? 5. Would appreciate an explanation on how authors decided on confounders to include in model and whether results differed in age-adjusted vs. multivariable models. 6. Might be interesting to see if associations differ by BMI 7. Discussion section could benefit from a more thorough explanation of findings from prior studies of other dietary scores (AHEI, aMED, DASH) and breast cancer risk; specifically focusing on stronger findings with ER- tumors. The results from the current study are not necessarily in line with results from many studies which suggest null relations between diet quality and breast cancer risk overall, although inverse associations have been observed for ER- subtypes. Authors should address why they believe that the diet score in this study was strongly associated with overall breast cancer risk in this study.

REVIEWER	Teresa Fung Simmons College, USA
REVIEW RETURNED	16-Oct-2016

GENERAL COMMENTS	This is a worthwhile analysis and the authors have a decent data set. The authors should take better advantage of their dataset to dig a little deeper in the analysis. The writing needs some English editing. It is not terrible, but many places can use some refinement. For example, in the first sentence of the introduction, "first" is not right word, the authors probably meant "most common". One major issue is that the methods needs more details, see specific

	comments below. Specific comments. Abstract -- Please add more information on the FAS-NPS DI. Add that analysis was multivariable adjusted. Introduction, reference 2 -- Is that specific to breast cancer? Please use a reference specific to breast cancer. Need some clarity on the 24-hour recall. In a 2-week period, three 24-hour recalls were done. How many sets of these three 24-hour recall was done in what period of time (e.g. 1 year). Later on, it was alluded that some participants has six 24-hour recall. Would that be from two 3-day sets or that it could be from 3 sets but each set has only 2 recalls? And over what period of time were the recalls conducted? Only in the first 2 years or throughout the follow-up period? The scoring algorithm of the FSA-NPS DI is clear. But more explanation is needed on how the 24-hour recalls were used to calculate them. Were the recalls averaged to give a mean 24-hour intake and then the scoring algorithm was applied? Or that for each recall, a FSA-NPD DI was computed and then averaged over all the recalls? It is unusual that a diet quality score would award points for unhealthy components and so a higher score represents an unhealthy diet. This makes comparison with other diet quality scores more cumbersome. For breast cancer cases, were all stages included? It would be worthwhile to exclude stage 4 cases as diet might have less impact in late stage diseases. Also, in situ tumors are not exactly invasive yet, therefore these should also be excluded for a cleaner analysis. Were all self-reported cases confirmed by medical records or other methods? More clarity is needed for when follow-up time began. There is the date of the baseline questionnaires were returned, and then the 24-recalls were done at different dates, with different participants having different numbers of 24-hour recall. Results -- please report the number of cases in each cancer stage. Since ER status is known, please separately analyze the data by ER status. The authors stated that results were similar for pre- and post menopausal cases, but that is apparent only for the results on 1-point increment of the diet score. It appears that despite some participants have quite a number of 24-hour recalls, these were assessed in the early part of follow-up but not throughout the entire follow-up period. Given the long follow-up, diet could have changed, therefore, this should be mentioned as a limitation.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1
Reviewer Name: Holly R. Harris
Institution and Country: Assistant Member, Fred Hutchinson Cancer Research Center, United States
Please state any competing interests or state 'None declared': None declared.

Please leave your comments for the authors below
This manuscript presents the evaluation of the FSA NPS DI, which measures the nutritional quality of the diet, and breast cancer risk. A thorough discussion of the use of different food indexes is included in the discussion section. Overall the study was well-designed and the paper is well-written, however there are a few ways in which the manuscript could be improved.
→ We thank Reviewer 1 for this positive comment.

Methods:
How was death ascertained? Please describe.
→ Death was ascertained through linkage to the national database on mortality for the French population (Inserm-CépiDC). This has been added in the manuscript, p7-8: "Information on death and cause of death was obtained through linkage to the national database on mortality of the French population [1]."

Were there any breast cancer cases in age <35 years? How does including women <35 years in the analysis influence the results? If women turned 35 during the course of the study were they included in the analysis?
→ In this study, we excluded women aged <35y at the time of their inclusion in the cohort (thus, women that turned 35 during the course of the study were not included). In the overall NutriNet-Santé cohort, participants are included from the age of 18, which may induce some noise in the cancer analyses since younger women are substantially less at risk of developing breast cancer [2] and since nutrition may have a lesser effect on breast cancers diagnosed in young women. This has been clarified in the manuscript, p8: "Women aged <35y at baseline (n=29,249) were excluded due to a very low susceptibility to develop breast cancer in these women [2] and a potentially limited influence of nutrition on breast cancers diagnosed in young women."
We have performed sensitivity analyses with the inclusion of women that were aged <35y at baseline but who turned 35 during the course of the study (contribution to the model from the moment when they turned 35). Similar results were observed: 568 cases/ 52,258 non-cases, HRQ5vs.Q1=1.47 (95%CI 1.08-2.02), P-trend=0.009; HRper 1-unit increment=1.06 (1.02-1.11), P=0.007.
555 breast cancers were diagnosed in women aged above 35 at the time of inclusion and 30 in women aged under 35 at the time of inclusion (13 were diagnosed at age 35-40 and 17 at age under 35).
Following Reviewer 1's suggestions, analyses have been performed without the exclusion of women aged under 35 at baseline. This has been added in the manuscript, p11: "Similar but weaker trends were also observed when women aged <35y at baseline were included in the analyses (585 cases/ 74,617 non-cases, HRQ5vs.Q1=1.17 (95%CI 0.83-1.64), P-trend=0.1; HRper 1-unit increment=1.05 (1.01-1.10), P=0.02)".

Results:
The authors describe the results as being similar in premenopausal and postmenopausal women but the HR for premenopausal women is much stronger (HR 2.46 vs 1.25) and the result for postmenopausal women (the group with the most power) is not significant. The difference in the point estimates should be more accurately described. In addition, a more formal test for heterogeneity between these estimates should be conducted and presented in the results.
→ We agree that results for pre- and post-menopausal women were similar in the direction but not in the strength of the association. Pre-menopausal women were more likely to score high on the FSA-

NPS DI: mean ± SD FSA-NPS DI was 6.3±2.3 in women pre-menopause (median:6.4, 25th-75th percentiles: 4.9-7.8) and 5.5±2.1 in women post-menopause (median:5.5, 25th-75th percentiles: 4.1-6.9), reflecting a lower nutritional quality of the diet in younger subjects, as expected. As observed in Table 1, women pre-menopause were more likely to pertain to the highest quintiles of FSA-NPS DI whereas women post-menopause were more likely to pertain to the lowest quintiles. This may explain the stronger effect observed in women pre-menopause.
The presentation of the results has been modified to better reflect the observed difference and now includes a formal test of interaction between menopausal status and FSA-NPS DI, p10 (and Table 2): "These associations were similarly observed in premenopausal women (HRQ5vs.Q1=2.46 (1.27-4.75), P-trend=0.004; HRper 1-unit increment=1.09 (1.01-1.18), P=0.03) and in postmenopausal women (HRQ5vs.Q1=1.25 (0.85-1.84), P-trend=0.09; HRper 1-unit increment=1.05 (1.00-1.11), P=0.06), although the associations seemed stronger for pre-menopausal women and only trends were observed for post-menopausal women (P-interaction=0.06)."
Some discussion has been added to the manuscript, p14: "In our study, although similar trends were observed in pre- and post-menopausal women for the association between the FSA-NPS DI and breast cancer risk, this association was nonetheless stronger in pre-menopausal women. This may be explained by the fact that women pre-menopause were more likely to score high on the FSA-NPS DI, thus resulting in a clearer/stronger association: mean±SD FSA-NPS DI was 6.3±2.3 in women pre-menopause (median:6.4, 25th-75th percentiles: 4.9-7.8) and 5.5±2.1 in women post-menopause (median:5.5, 25th-75th percentiles: 4.1-6.9)."

Results are not presented by hormone receptor status due to lack of power. However, given the differing results by menopausal status is would also be interesting to see the results by hormone receptor status, at least by ER+/PR+ and ER-/PR-, with the caveat of limited power.
→ Following Reviewer 1, 2 and 3's suggestions, results are now presented by hormone receptor status, p11 and Supplementary file 3: "Information regarding hormone receptor status was not available for all cases (ER status: 361 cases, PR status: 362 cases, ER/PR status: 361 cases). Significant direct associations between the FSA-NPS DI and breast cancer risk were observed for breast cancer types PR- (102 cases/ 46,762 non-cases) and ER+/PR- (49 cases/ 46,815 non-cases). For ER+ tumours, the linear trend was not statistically significant (P=0.07, 307 cases/46,557 non-cases) but compared to women in the lowest quintile of FSA-NPS DI, those with higher scores had an increased breast cancer risk (e.g. HR Q5vs.Q1=1.60 (1.04-1.46)). Associations were non-significant for the other hormone receptor status (Supplementary file 3). However, these exploratory findings should be considered with caution due to limited statistical power for analyses by cancer sub-types."
Partial availability of the data regarding hormonal status and limited statistical power in these analyses are discussed, p13: "In our study, information regarding hormonal receptor status of the tumors was only partially available and the statistical power was limited in the analyses (Supplementary file 3), thus preventing to derive firm conclusions."

Discussion:
A discussion of stronger results were observed for premenopausal compared to postmenopausal women is needed.
→ As stated above, some discussion has been added to the manuscript, p14: "In our study, although similar trends were observed in pre- and post-menopausal women for the association between the FSA-NPS DI and breast cancer risk, this association was nonetheless stronger in pre-menopausal women. This may be explained by the fact that women pre-menopause were more likely to score high on the FSA-NPS DI, thus resulting in a clearer/stronger association: mean±SD FSA-NPS DI was 6.3±2.3 in women pre-menopause (median:6.4, 25th-75th percentiles: 4.9-7.8) and 5.5±2.1 in women post-menopause (median:5.5, 25th-75th percentiles: 4.1-6.9)."

The median follow-up time in this study is 4 years which is shorter than many previous studies that have examined dietary factors and breast cancer. The paper could be improved with a discussion of

how the timing of the diet in relation to breast cancer diagnosis could influence the results. Given that in this study dietary intake is relatively close to breast cancer diagnosis and cancer may take years to develop does this indicate more a promoter effect (in contrast to being an initiator) of diet on breast cancer? How does this compare to what has been observed with previous studies that have examined dietary indices and breast cancer risk?
→ Some discussion has been added p15: “Finally, as usually done in nutritional epidemiology, dietary intakes were estimated based on averaged intakes from all 24h-dietary records collected over the first two years of follow-up. Although diet may change over time, it is usually hypothesized that this estimation reflects general eating behavior throughout the adult life [3]. This very classical method allowed us to obtain a reliable estimation of usual dietary intakes, while respecting the prospective design (i.e. estimation of usual dietary intakes prior to cancer diagnosis). Indeed, breast cancer is a disease with relatively long latency so that the involvement of nutritional factors is supposed to be based on long-term processes. Thus, it is important to guarantee sufficient delay between nutritional exposure and cancer outcome. This is why we tested a model (sensitivity analysis) where cancer cases diagnosed during the first year of follow-up were excluded (similar results). In our study, although the follow-up time was appropriate to perform etiological analyses, it did not necessarily guarantee this sufficient delay. Hence, our estimation of usual dietary intakes may reflect dietary protective and risk factors that may have played a role in the first steps of carcinogenesis (initiation) but also later in the carcinogenic process (progression). Nonetheless, previous studies with longer follow-up observed associations between diet and breast cancer risk, suggesting that nutritional factors could play a role in cancer initiation and not only in cancer progression [4-9].”

Table 1:
A few key foods/foods groups/etc could be included in Table 1 (i.e. servings of red or processed meat, fiber, vegetable intake, etc).
→ Table 1 now displays intakes for several nutrients (lipids, carbohydrates, proteins, fibers) and food groups (fruit, vegetable, legume, red meat, processed meat, poultry, fish, dairy) overall and by quintiles of the FSA-NPS DI. As expected, women with a lower FSA-NPS DI (diet of higher nutritional quality) had overall healthier food intakes: higher intakes of fiber, fruits, vegetables, legume, fish and lower intakes of red and processed meat and lipids. This has been added in the manuscript, p10.

BMI does not yet exhibit a clear trend (Q5 has the second highest mean BMI with Q1 the highest) yet the p for trend is very significant. How is this explained?
→ In the NutriNet-Santé study, given the large sample size, results from descriptive analyses generally happen to be significant even in the case of small differences.
We have added in Table 1 the distribution of BMI (normal-weight, overweight, obese) by quintiles of the FSA-NPS DI showing somehow clearer trends.

Table 2:
It would be helpful to presented the age-adjusted models as well.
→ Age-adjusted analyses are now presented in Table 2.

Reviewer 2
Reviewer Name: Kelly Hirko
Institution and Country: Michigan State University College of Human Medicine, USA
Please state any competing interests or state ‘None declared’: NONE DECLARED

Please leave your comments for the authors below
1. It would be interesting to evaluate which component(s) of the FSA NPS score were most strongly associated with breast cancer risk
→ Unlike other a priori scores, components of the FSA-NPS DI cannot be studied separately. Indeed,

the FSA-NPS DI is first calculated at the food level (FSA-NPS) and then aggregated at the individual level. This is at the food level that the score summarizes several components. Furthermore, the calculation of the FSA-NPS score (see Supplemental file 1) is based on thresholds and is conditional (e.g. points for proteins are taken into account only if the food respects some conditions).

2. Although authors likely don't have power to look at ER- tumors, it might be interesting to conduct sensitivity analyses only among the ER+ to see if the associations appear weaker, given the scientific evidence that nutritional factors may be more important in ER- breast cancer.
→ Following Reviewer 1, 2 and 3's suggestions, results are now presented by hormone receptor status, p11 and Supplementary file 3: "Information regarding hormone receptor status was not available for all cases (ER status: 361 cases, PR status: 362 cases, ER/PR status: 361 cases). Significant direct associations between the FSA-NPS DI and breast cancer risk were observed for breast cancer types PR- (102 cases/ 46,762 non-cases) and ER+/PR- (49 cases/ 46,815 non-cases). For ER+ tumours, the linear trend was not statistically significant (P=0.07, 307 cases/46,557 non-cases) but compared to women in the lowest quintile of FSA-NPS DI, those with higher scores had an increased breast cancer risk (e.g. HR Q5vs.Q1=1.60 (1.04-1.46)). Associations were non-significant for the other hormone receptor status (Supplementary file 3). However, these exploratory findings should be considered with caution due to limited statistical power for analyses by cancer sub-types." Partial availability of the data regarding hormonal status and limited statistical power in these analyses are discussed, p13: "In our study, information regarding hormonal receptor status of the tumors was only partially available and the statistical power was limited in the analyses (Supplementary file 3), thus preventing to derive firm conclusions."

3. What do authors mean by 'women with a null follow-up'? Please clarify
→ In the NutriNet-Santé study, the beginning of the follow-up was determined as the date of last completed baseline questionnaire and the end of follow-up was determined as the date of cancer diagnosis, the date of last completed questionnaire, the date of death or August 2015, whichever occurred first. For a small proportion of women, the last completed questionnaire was the last completed baseline questionnaire, resulting in a null follow-up (end of follow-up equals beginning of follow-up). This was clarified, p8: "Women with a null follow-up were also excluded from the analyses (i.e. women for whom baseline questionnaires were the last completed questionnaires, n=921)"

4. How many women in the NutriNet-Sante cohort were excluded from this study because they didn't provide at least 3 24-h dietary records? Did those excluded differ from included participants based on demographic characteristics related to breast cancer risk?
→ In the NutriNet-Santé study, 15,918 women (non-null follow-up) provided less than three 24h-dietary records over their first two years of follow-up. A comparison of women that provided at least three 24h-dietary records and women that did not has been added in the manuscript, p10: "Compared to women that provided at least three 24h-dietary records over their first two years of follow-up, women that did not (15,918 women with a non-null follow-up) were younger, pre-menopause, were more likely to be overweight/obese, to smoke, to practice physical activity and were less likely to have a family history of cancer or to take a hormonal treatment for menopause [data not tabulated]."

5. Would appreciate an explanation on how authors decided on confounders to include in model and whether results differed in age-adjusted vs. multivariable models.
→ The confounders included in our models were classic risk factors for breast cancer (as done in our previous study of the association between the FSA-NPS DI and cancer risk [10]). This has been clarified in the manuscript, p8: "Models were adjusted for classic risk factors for breast cancer". Besides, age-adjusted analyses are now presented in Table 2. Results were overall similar, with strengthened associations (especially for post-menopausal women for whom results were statistically significant in the age-adjusted analyses).

6. Might be interesting to see if associations differ by BMI
→ Following Reviewer 2's suggestion, we have tested an interaction between BMI and the FSA-NPS DI, p8: "Interaction analysis was conducted between BMI and the FSA-NPS DI and stratified analyses were performed according to overweight status (BMI < vs. ≥25kg/m2)."
Results of the stratified analyses are shown p10-11: "Analyses performed according to overweight status showed that associations tended to be stronger in non-overweight women (368 cases/ 31,401 non-cases, HRQ5vs.Q1=1.97 (95%CI 1.31-2.96), P-trend=0.0007; HRper 1-unit increment=1.09 (1.03-1.15), P=0.003) compared to overweight/obese women (187 cases/14,908 non-cases, HRQ5vs.Q1=1.02 (95%CI 0.61-1.73), P-trend=0.6; HRper 1-unit increment=1.03 (0.95-1.11), P=0.5), but the interaction was not statistically significant (P=0.07)."

7. Discussion section could benefit from a more thorough explanation of findings from prior studies of other dietary scores (AHEI, aMED, DASH) and breast cancer risk; specifically focusing on stronger findings with ER- tumors. The results from the current study are not necessarily in line with results from many studies which suggest null relations between diet quality and breast cancer risk overall, although inverse associations have been observed for ER- subtypes. Authors should address why they believe that the diet score in this study was strongly associated with overall breast cancer risk in this study.
→ In our study, it was not possible to properly investigate ER- breast cancers due to a lack of statistical power (even though results for ER- tumors have been provided following the reviewers' suggestions). The discussion regarding findings from prior studies has been strengthened to include differences observed according to hormonal receptor status of the tumors, p13: "In these studies, differences according to hormonal receptor status of the tumors have been suggested, with inconsistent results. Indeed, inverse associations between a "healthier" diet and breast cancer risk were particularly observed in ER- type (AHEI, RFS, aMed) [5], ER-/PR+ type (Mediterranean diet score) [11], and ER-/PR-/HER2+ type (DASH) [7], but also with ER+/PR+ type (WCRF/AICR adherence score) [6] and ER+/PR- type (''healthy/Mediterranean'' pattern) [4]. In our study, information regarding hormonal receptor status of the tumors was only partially available and the statistical power was limited in the analyses (Supplementary file 3), thus preventing to derive firm conclusions."
The FSA-NPS DI was not the only dietary quality score for which a decreased breast cancer risk was observed, since a decreased breast cancer risk was also observed in studies investigating dietary scores measuring the adherence to cancer-specific nutritional recommendations (WCRF/AICR adherence score, ACS cancer prevention guidelines score) [6;8;9]. However, the comparison between the FSA-NPS DI and other dietary scores is not straightforward, as stated in the manuscript, p13-14: "Overall, these studies involving a priori scores provided interesting insights into the relationships between nutrition and breast cancer risk. Although these a priori scores and the FSA-NPS DI included similar nutritional components, the approaches differed, making the comparison between our study and previous findings not straightforward (even though our results were in line with those obtained with scores measuring the adherence to cancer-specific nutritional recommendations [6;8;9]). The FSA-NPS DI is not primarily built at the individual level but is rather derived from a nutrient profiling system at the food level (FSA-NPS) thus taking into account the nutritional quality of each food/beverage consumed and not only of the overall diet or overall consumption of food groups. In addition, the objective behind the FSA-NPS DI construction was not to obtain the best predictive score but to specifically test its association with breast cancer risk, as the FSA-NPS is envisioned to serve as a basis for food labelling in the framework of public health policies in several countries such as France and Australia."
In addition, in the NutriNet-Santé cohort, as stated p14, "Strengths of this study pertained to its prospective design, its large sample size, and the assessment of usual dietary intakes using repeated 24h-dietary records based on a recent food composition database with a large choice of items (>3300). The latter allowed a better insight into the food products consumed and their intrinsic nutritional quality compared to studies that used a food frequency questionnaire (more aggregated

food items).”

Reviewer 3
Reviewer Name: Teresa Fung
Institution and Country: Simmons College, USA
Please state any competing interests or state ‘None declared’: None declared.

Please leave your comments for the authors below
This is a worthwhile analysis and the authors have a decent data set. The authors should take better advantage of their dataset to dig a little deeper in the analysis.
→ We thank Reviewer 3 for this positive comment.

The writing needs some English editing. It is not terrible, but many places can use some refinement. For example, in the first sentence of the introduction, "first" is not right word, the authors probably meant "most common".
→ Some English editing have been made throughout the manuscript to refine the writing.

One major issue is that the methods needs more details, see specific comments below.
Specific comments.
Abstract -- Please add more information on the FAS-NPS DI. Add that analysis was multivariable adjusted.
→ The abstract now displays more information regarding the FSA-NPS DI, p2: “The FSA-NPS is calculated for each food/beverage based on the amount per 100g of energy, total sugar, saturated fatty acid, sodium, dietary fibers, proteins, and % of fruits and vegetables.”, and states that analyses were multivariable adjusted: “Associations between individual FSA-NPS DI and breast cancer risk were characterized by multivariable-adjusted Cox proportional hazard models.”

Introduction, reference 2 -- Is that specific to breast cancer? Please use a reference specific to breast cancer.
→ Reference 2 [12] refers to an online table released by the World Cancer Research Fund on cancer preventability estimates for diet, nutrition, body fatness, and physical activity for 4 countries (USA, UK, Brazil, China). In this table, specific estimates for breast cancer are included.

Need some clarity on the 24-hour recall. In a 2-week period, three 24-hour recalls were done. How many sets of these three 24-hour recall was done in what period of time (e.g. 1 year). Later on, it was alluded that some participants has six 24-hour recall. Would that be from two 3-day sets or that it could be from 3 sets but each set has only 2 recalls? And over what period of time were the recalls conducted? Only in the first 2 years or throughout the follow-up period?
→ Some precisions have been added regarding the 24-dietary records to improve clarity, p6: “Dietary intakes were assessed at baseline and every six months through series of three non-consecutive validated web-based 24h-dietary records, randomly assigned over a 2-week period (2 weekdays and 1 weekend day) [13-15]. Thus, over the first two years of follow-up, up to five series of three 24h-dietary records could have been completed. To be considered as valid, a series must have included at least two out of three 24h dietary records. Participants used a dedicated interface of the study website to declare all foods and beverages consumed during a 24h-period: three main meals (breakfast, lunch, dinner) or any other eating occasion. Portion sizes were estimated using validated photographs [16]. Mean daily energy, alcohol and nutrient intakes were estimated using a published French food composition table (>3300 items) [17] and a weighting for week days and week-end days.” Therefore, women with at least three 24h-dietary records must have provided at minimum 1 series of 3 24h-dietary records or 2 series of 2 24h-dietary records. Sensitivity analyses were performed including only women who provided at least six 24h-dietary records during their first two years of

follow-up: i.e. at minimum 2 series of 3 24h-dietary records, 3 series of 2 24h-dietary records, 2 series of 2 + 1 series of 3 24h-dietary records etc.

The scoring algorithm of the FSA-NPS DI is clear. But more explanation is needed on how the 24-hour recalls were used to calculate them. Were the recalls averaged to give a mean 24-hour intake and then the scoring algorithm was applied? Or that for each recall, a FSA-NPD DI was computed and then averaged over all the recalls?
→ The first solution is right. For each participant, all 24h-dietary records from their first two years of follow-up were averaged to give a mean 24-hour intake for each food/beverage consumed and the corresponding mean 24h-energy intake from this food/beverage. The FSA-NPS DI is then calculated for each participant as an energy-weighted mean of the FSA-NPS scores of all foods/beverages consumed. This has been clarified in the manuscript, p7: “In a second step, the FSA-NPS DI was computed at the individual level using arithmetic energy-weighted means with the following equation [18], in which FSi represents the food (or beverage) score, and Ei represents energy intake from this food or beverage (all 24h-dietary records from the first two years of follow-up were averaged to a mean 24-hour energy intake from this food/beverage)”.

It is unusual that a diet quality score would award points for unhealthy components and so a higher score represents an unhealthy diet. This makes comparison with other diet quality scores more cumbersome.
→ Indeed, the FSA-NPS was originally built and published with points awarded for unhealthy components [19;20] and the deriving FSA-NPS DI was used as such in several studies ever since [10;18;21-24]. Although this construction may be unusual compared to other dietary quality scores, comparison with these other scores is already not straightforward since the FSA-NPS DI is based on a nutrient profiling system at the food level (FSA-NPS) from which is then derived a score at the individual level. The objective of this study was not to find the best dietary quality score that would be associated to a decrease in breast cancer risk. Our objective was rather to assess the relevance of the implementation of a system, the FSA-NPS, to grade the nutritional quality of foods/beverages for the purpose of front-of-pack labelling. Although the scale is reversed, the important point is that, overall, the same components are considered as ‘healthy’ or ‘less healthy’ in the FSA-NPS DI and the other dietary quality scores.

For breast cancer cases, were all stages included? It would be worthwhile to exclude stage 4 cases as diet might have less impact in late stage diseases.
→ In our analyses, all breast cancer cases were included, regardless of the stage. Information regarding breast cancer stage was not coded in the database so that it was not possible to perform analyses excluding stage 4. This has been added as a limitation, p15: “Next, information regarding cancer stage was not available.”

Also, in situ tumors are not exactly invasive yet, therefore these should also be excluded for a cleaner analysis.
→ We performed sensitivity analyses excluding breast cancers in situ and obtained similar results, p11: “Similar results were observed […] when analyses were restricted to invasive breast cancers (387 cases/46,309 non-cases; HRQ5vs.Q1=1.51 (1.03-2.22), P-trend=0.01; HRper 1-unit increment=1.06 (1.01-1.12), P=0.03))”
Since the association between the FSA-NPS DI and in situ breast cancer was also of interest, and since information regarding the in situ/invasive type of the tumors was only available for 463/555 cases, we chose not to exclude in situ breast cancers in our main analyses.

Were all self-reported cases confirmed by medical records or other methods?
→ As stated in the methods, p7: “Participants self-declared health events through the yearly health status questionnaire, through a specific check-up questionnaire for health events (every three

months) or at any time through a dedicated interface on the study website. Following this declaration, participants are invited to send their medical records (diagnosis, hospitalization, etc.) and, if necessary, the study physicians contact the participants' treating physician or the medical structures to collect additional information. Then, data are reviewed by an independent physician expert committee which validates all major health events (such as cancers)."

More clarity is needed for when follow-up time began. There is the date of the baseline questionnaires were returned, and then the 24-recalls were done at different dates, with different participants having different numbers of 24-hour recall.
→ Follow-up of participants began when participants answered their last baseline questionnaire. The date of completion of the last baseline questionnaire is thus used as inclusion date. This has been clarified in the manuscript p6.

Results -- please report the number of cases in each cancer stage. Since ER status is known, please separately analyze the data by ER status.
→ Information regarding breast cancer stage was not coded in the database. This has been added as a limitation, p15: "information regarding cancer stage was not available".
Following Reviewer 1, 2 and 3's suggestions, results are now presented by hormone receptor status, p11 and Supplementary file 3: "Information regarding hormone receptor status was not available for all cases (ER status: 361 cases, PR status: 362 cases, ER/PR status: 361 cases). Significant direct associations between the FSA-NPS DI and breast cancer risk were observed for breast cancer types PR- (102 cases/ 46,762 non-cases) and ER+/PR- (49 cases/ 46,815 non-cases). For ER+ tumours, the linear trend was not statistically significant (P=0.07, 307 cases/46,557 non-cases) but compared to women in the lowest quintile of FSA-NPS DI, those with higher scores had an increased breast cancer risk (e.g. HR Q5vs.Q1=1.60 (1.04-1.46)). Associations were non-significant for the other hormone receptor status (Supplementary file 3). However, these exploratory findings should be considered with caution due to limited statistical power for analyses by cancer sub-types."
Partial availability of the data regarding hormonal status and limited statistical power in these analyses are discussed, p13: "In our study, information regarding hormonal receptor status of the tumors was only partially available and the statistical power was limited in the analyses (Supplementary file 3), thus preventing to derive firm conclusions."

The authors stated that results were similar for pre- and post menopausal cases, but that is apparent only for the results on 1-point increment of the diet score.
→ We agree that results for pre- and post-menopausal women were similar in the direction but not in the strength of the association. Pre-menopausal women were more likely to score high on the FSA-NPS DI: mean±SD FSA-NPS DI was 6.3±2.3 in women pre-menopause (median:6.4, 25th-75th percentiles: 4.9-7.8) and 5.5±2.1 in women post-menopause (median:5.5, 25th-75th percentiles: 4.1-6.9), reflecting lower nutritional quality of the diet in younger subjects, as expected. As observed in Table 1, women pre-menopause were more likely to pertain to the highest quintiles of FSA-NPS DI whereas women post-menopause were more likely to pertain to the lowest quintiles. This may explain the stronger effect observed in women pre-menopause.
The presentation of the results has been modified to better reflect the observed difference and now includes a formal test of interaction between menopausal status and FSA-NPS DI, p10 (and Table 2): "These associations were similarly observed in premenopausal women (HRQ5vs.Q1=2.46 (1.27-4.75), P-trend=0.004; HRper 1-unit increment=1.09 (1.01-1.18), P=0.03) and in postmenopausal women (HRQ5vs.Q1=1.25 (0.85-1.84), P-trend=0.09; HRper 1-unit increment=1.05 (1.00-1.11), P=0.06), although the associations seemed stronger for pre-menopausal women and only trends were observed for post-menopausal women (P-interaction=0.06)."
Some discussion has been added to the manuscript, p14: "In our study, although similar trends were observed in pre- and post-menopausal women for the association between the FSA-NPS DI and breast cancer risk, this association was nonetheless stronger in pre-menopausal women. This may be

explained by the fact that women pre-menopause were more likely to score high on the FSA-NPS DI, thus resulting in a clearer/stronger association: mean±SD FSA-NPS DI was 6.3±2.3 in women pre-menopause (median:6.4, 25th-75th percentiles: 4.9-7.8) and 5.5±2.1 in women post-menopause (median:5.5, 25th-75th percentiles: 4.1-6.9).”

It appears that despite some participants have quite a number of 24-hour recalls, these were assessed in the early part of follow-up but not throughout the entire follow-up period. Given the long follow-up, diet could have changed, therefore, this should be mentioned as a limitation.
→ Some discussion has been added p15: “Finally, as usually done in nutritional epidemiology, dietary intakes were estimated based on averaged intakes from all 24h-dietary records collected over the first two years of follow-up. Although diet may change over time, it is usually hypothesized that this estimation reflects general eating behavior throughout the adult life [3]. This very classical method allowed us to obtain a reliable estimation of usual dietary intakes, while respecting the prospective design (i.e. estimation of usual dietary intakes prior to cancer diagnosis). Indeed, breast cancer is a disease with relatively long latency so that the involvement of nutritional factors is supposed to be based on long-term processes. Thus, it is important to guarantee sufficient delay between nutritional exposure and cancer outcome. This is why we tested a model (sensitivity analysis) where cancer cases diagnosed during the first year of follow-up were excluded (similar results).”

References

(1) Inserm. CépiDC. Interrogation des données sur les causes de décès de 1979 à 2013. Inserm [2016 Available from: URL:http://www.cepidc.inserm.fr/inserm/html/index2.htm
(2) INCa. Incidence et mortalité estimées par classe d'âge et par localisation cancéreuse en 2012. 2016.
(3) Willett WC. Nutritional Epidemiology, 2nd ed. New York: Oxford University Press; 1998.
(4) Cottet V, Touvier M, Fournier A, Touillaud MS, Lafay L, Clavel-Chapelon F et al. Postmenopausal breast cancer risk and dietary patterns in the E3N-EPIC prospective cohort study. Am J Epidemiol 2009; 170(10):1257-1267.
(5) Fung TT, Hu FB, McCullough ML, Newby PK, Willett WC, Holmes MD. Diet quality is associated with the risk of estrogen receptor-negative breast cancer in postmenopausal women. J Nutr 2006; 136(2):466-472.
(6) Harris HR, Bergkvist L, Wolk A. Adherence to the World Cancer Research Fund/American Institute for Cancer Research recommendations and breast cancer risk. Int J Cancer 2016; 138(11):2657-2664.
(7) Hirko KA, Willett WC, Hankinson SE, Rosner BA, Beck AH, Tamimi RM et al. Healthy dietary patterns and risk of breast cancer by molecular subtype. Breast Cancer Res Treat 2016; 155(3):579-588.
(8) Nomura SJ, Inoue-Choi M, Lazovich D, Robien K. WCRF/AICR recommendation adherence and breast cancer incidence among postmenopausal women with and without non-modifiable risk factors. Int J Cancer 2016; 138(11):2602-2615.
(9) Thomson CA, McCullough ML, Wertheim BC, Chlebowski RT, Martinez ME, Stefanick ML et al. Nutrition and physical activity cancer prevention guidelines, cancer risk, and mortality in the women's health initiative. Cancer Prev Res (Phila) 2014; 7(1):42-53.
(10) Donnenfeld M, Julia C, Kesse-Guyot E, Mejean C, Ducrot P, Peneau S et al. Prospective association between cancer risk and an individual dietary index based on the British Food Standards Agency Nutrient Profiling System. Br J Nutr 2015; 114(10):1702-1710.
(11) Schwingshackl L, Hoffmann G. Adherence to Mediterranean diet and risk of cancer: an updated systematic review and meta-analysis of observational studies. Cancer Med 2015; 4(12):1933-1947.
(12) WCRF/AICR. Cancer preventability estimates for food, nutrition, body fatness, and physical activity. WCRF/AICR [2016 Available from: URL:http://www.wcrf.org/int/cancer-facts-figures/preventability-estimates/cancer-preventability-estimates-diet-nutrition

(13) Lassale C, Castetbon K, Laporte F, Deschamps V, Vernay M, Camilleri GM et al. Correlations between Fruit, Vegetables, Fish, Vitamins, and Fatty Acids Estimated by Web-Based Nonconsecutive Dietary Records and Respective Biomarkers of Nutritional Status. J Acad Nutr Diet 2016; 116(3):427-438.
(14) Lassale C, Castetbon K, Laporte F, Camilleri GM, Deschamps V, Vernay M et al. Validation of a Web-based, self-administered, non-consecutive-day dietary record tool against urinary biomarkers. Br J Nutr 2015; 113(6):953-962.
(15) Touvier M, Kesse-Guyot E, Mejean C, Pollet C, Malon A, Castetbon K et al. Comparison between an interactive web-based self-administered 24 h dietary record and an interview by a dietitian for large-scale epidemiological studies. Br J Nutr 2011; 105(7):1055-1064.
(16) Le Moullec N, Deheeger M, Preziosi P, Montero P, Valeix P, Rolland-Cachera M et al. Validation du manuel photo utilisé pour l'enquête alimentaire de l'étude SU.VI.MAX. [Validation of the food portion size booklet used in the SU.VI.MAX study] (in French). Cah Nutr Diet 1996; 31:158-164.
(17) Arnault N, Caillot L, Castetbon K, Coronel S, Deschamps V, Fezeu L et al. Table de composition des aliments, étude NutriNet-Santé. [Food composition table, NutriNet-Santé study] (in French). Paris: Les éditions INSERM/Economica; 2013.
(18) Julia C, Touvier M, Mejean C, Ducrot P, Peneau S, Hercberg S et al. Development and validation of an individual dietary index based on the British Food Standard Agency nutrient profiling system in a French context. J Nutr 2014; 144(12):2009-2017.
(19) Arambepola C, Scarborough P, Rayner M. Validating a nutrient profile model. Public Health Nutr 2008; 11(4):371-378.
(20) Rayner M, Scarborough P, Stockley P, Boxer A. Nutrient profiles: Development of Final Model. Final Report [online]. London, FSA . 2005.
(21) Adriouch S, Julia C, Kesse-Guyot E, Mejean C, Ducrot P, Peneau S et al. Prospective association between a dietary quality index based on a nutrient profiling system and cardiovascular disease risk. Eur J Prev Cardiol 2016.
(22) Julia C, Ducrot P, Lassale C, Fezeu L, Mejean C, Peneau S et al. Prospective associations between a dietary index based on the British Food Standard Agency nutrient profiling system and 13-year weight gain in the SU.VI.MAX cohort. Prev Med 2015; 81:189-194.
(23) Julia C, Mejean C, Touvier M, Peneau S, Lassale C, Ducrot P et al. Validation of the FSA nutrient profiling system dietary index in French adults-findings from SUVIMAX study. Eur J Nutr 2015.
(24) Julia C, Fezeu LK, Ducrot P, Mejean C, Peneau S, Touvier M et al. The Nutrient Profile of Foods Consumed Using the British Food Standards Agency Nutrient Profiling System Is Associated with Metabolic Syndrome in the SU.VI.MAX Cohort. J Nutr 2015; 145(10):2355-2361.

VERSION 2 – REVIEW

REVIEWER	Holly R. Harris Fred Hutchinson Cancer Research Center, U.S.A.
REVIEW RETURNED	07-Jan-2017

GENERAL COMMENTS	The authors have adequately addressed all my previous comments.

REVIEWER	Kelly Hirko Michigan State University College of Human Medicine, USA
REVIEW RETURNED	15-Jan-2017

GENERAL COMMENTS	The authors adequately addressed my prior concerns and I have no further comments

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1
The authors have adequately addressed all my previous comments.
=> We thank Reviewer 1 for this positive comment.

Reviewer: 2
The authors adequately addressed my prior concerns and I have no further comments
=> We thank Reviewer 2 for this positive comment.

References

(1) Julia C, Ducrot P, Lassale C, Fezeu L, Mejean C, Peneau S et al. Prospective associations between a dietary index based on the British Food Standard Agency nutrient profiling system and 13-year weight gain in the SU.VI.MAX cohort. Prev Med 2015; 81:189-194.
(2) Julia C, Fezeu LK, Ducrot P, Mejean C, Peneau S, Touvier M et al. The Nutrient Profile of Foods Consumed Using the British Food Standards Agency Nutrient Profiling System Is Associated with Metabolic Syndrome in the SU.VI.MAX Cohort. J Nutr 2015; 145(10):2355-2361.
(3) Adriouch S, Julia C, Kesse-Guyot E, Mejean C, Ducrot P, Peneau S et al. Prospective association between a dietary quality index based on a nutrient profiling system and cardiovascular disease risk. Eur J Prev Cardiol 2016.
(4) Donnenfeld M, Julia C, Kesse-Guyot E, Mejean C, Ducrot P, Peneau S et al. Prospective association between cancer risk and an individual dietary index based on the British Food Standards Agency Nutrient Profiling System. Br J Nutr 2015; 114(10):1702-1710.